2019

Structural Characterization of Black Widow Spider Dragline Silk Proteins CRP1 and CRP4

Mikayla Shanafelt
University of the Pacific, shanafeltm1@gmail.com

STRUCTURAL CHARATERIZATION OF BLACK WIDOW SPIDER SILK PROTEINS CRP1 AND CRP4

By

Mikayla Shanafelt

A Thesis Submitted to the

Graduate School

In Partial Fulfillment of the

Requirements for the Degree of

Master of Science

College of the Pacific
Biological Sciences

University of the Pacific
Stockton, California

2019

STRUCTURAL CHARATERIZATION OF BLACK WIDOW SPIDER SILK PROTEINS CRP1 AND CRP4

By

Mikayla Shanafelt

APPROVED BY:

Thesis Advisor: Craig A. Vierra, Ph.D.

Committee Member: Jane Khudyakov, Ph.D.

Committee Member: Geoff Lin-Cereghino, Ph.D.

Department Chair: Eric Thomas, Ph.D.

Dean of Graduate School: Thomas Naehr, Ph.D.

DEDICATION

For my Poppa Dean (August 19, 1942 – August 15, 2016). My work here is dedicated to you. Your love and support continue to push me to become all that I am. Lucky for me, Rocky Mountain Killer Quail are not native to Stockton, but if they were, I would have known what to do thanks to you. I love you more than words can describe, and I miss you tremendously. I can only hope to make you proud.

To my Papa Larry, my life at Pacific would not have existed without you. I am forever grateful for the opportunity and encouragement you provided for me.

And to my Mom, Alexa, my mentor, my cheering squad, my forever best friend. I love you and am incredibly thankful for everything you have done to lead me reach this moment.

ACKNOWLEDGEMENTS

This work was made possible by the support of my mentors, family, and friends. I am forever indebted to Dr. Craig Vierra for his continuous encouragement and guidance throughout my undergraduate and graduate career. Dr. V, working with you has been an honor and a pleasure. You have provided me a platform to become the best person, student, and scientist I can be, and I am truly thankful for everything you have done to forward my success.

Thank you to my thesis committee, Dr. Jane Khudyakov, and Dr. Geoff Lin-Cereghino, for your time and consideration of my research.

My gratitude also goes to Dr. Tara Thiemann, the first professor I ever attended a lecture from at Pacific. You were one of the major factors in leading me to decide that the University of the Pacific was where I wanted to start my undergraduate career. Your lecture was everything I was looking for in furthering my interest in biology and I walked out of your class that day knowing that I wanted to study at Pacific with professors like you.

My family, Pat, Alexa, Hailey, and Jack, thank you for always believing in me, for pushing me to limits I didn't know possible, and for listening to me explain all of my research (even if it wasn't quite as interesting to you as it was to me).

Jared, I would not be the person I am today without you and your amazing mind in my life. You have been my best friend in life and in science, never doubting my abilities and constantly challenging me to learn and achieve more. Thank you for being there for me without fail.

To my friends, PEAK, Taylor, Anish, thank you for all of the support you have provided me over the years and for making my life at Pacific full of adventures.

Structural Characterization of Black Widow Spider Dragline Silk Proteins CRP1 and CRP4

Abstract

By Mikayla Shanafelt

University of the Pacific
2019

Spider dragline silk is a biomaterial with outstanding material properties, possessing high-tensile strength and toughness. In nature, dragline silk serves a central role during spider locomotion and web construction. Today, scientists are racing to elucidate the molecular machinery governing silk extrusion, attempting to translate this knowledge into a mimicry process in the laboratory to create synthetic fibers for a wide range of different applications. During extrusion, it has been established that biochemical and mechanical forces govern spidroin folding, aggregation, and assembly. In black widow spiders, at least 7 different proteins have been identified as constituents of dragline silk fibers. These represent the major ampullate spidroins, MaSp1 and MaSp2, and several low-molecular weight cysteine-rich protein (CRP) family members that include CRP1, CRP2, and CRP4. Molecular modeling studies have predicted that CRPs contain a cystine slipknot motif. To advance scientific knowledge regarding CRP function, we expressed and purified recombinant CRP1 and CRP4 from bacteria and investigated their secondary structure using circular dichroism (CD) under different chemical and physical conditions. We demonstrate by far-UV CD spectroscopy that CRP1 and CRP4 contain similar secondary structural characteristics in solution, displaying substantial amounts of random coil conformation, followed by lower levels of beta sheet, alpha helical and beta-turn

structure. Additionally, we show that native structures of CRP1 and CRP4 have high thermal stability and are resistant to conformational changes under acidic pH conditions. Taken together, the chemical and thermal stability of CRP1 and CRP4 are experimentally consistent with biochemical properties of cystine slipknot proteins.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

LIST OF ABBREVIATIONS

Ala (A): Alanine

BME: β-mercaptoethanol

CD: Circular dichroism

CRP: Cysteine Rich Proteins

CTD/NTD: C-terminal domain/ N-terminal domain

Gly (G): Glycine

kDa: Kilodalton

MA: Major ampullate gland

MaSp: Major ampullate spidroin

μM, mM, M: Micromolar, millimolar, molar

μL, mL, L: Microliter, milliliter, liter

MS/MS: Mass spectrometry/mass spectrometry

Ni-NTA: Nickel-nitrilotriacetic acid

nm: Nanometer

PCR: Polymerase chain reaction

Pro (P): Proline

SDS-PAGE: Sodium dodecyl sulfate-polyacrylamide gel electrophoresis

Spidroin: Spider fibroin

SUMO: Small ubiquitin modifier

UV: Ultra Violet

X: Wobble Amino Acid

Chapter 1: Introduction

Spider silk has extraordinary mechanical properties, including high tensile strength and toughness.[1, 2] Spiders are capable of manufacturing 6-7 distinct fibers types, each finely tuned to perform a diverse range of biological functions.[3, 4] The most widely studied silk type in the scientific community is dragline silk. Dragline silk, also referred to as the "safety line" of the spider, is extruded from the major ampullate (MA) gland, a biofactory residing in the abdomen of the spider. Histological analyses have shown that the MA gland consists of 3 regions: a tail, ampulla, and spinning duct.[5, 6] Because dragline silk represents a high performance fiber that outperforms a vast majority of natural and man-made materials used in engineering, much effort has been focused on elucidating the constituents of dragline silk as well as the molecular mechanisms that control the pathway of silk extrusion.[6, 7]

Initially, dragline silk was described as a two-protein fiber, consisting of the spidroin family members MaSp1 and MaSp2.[8, 9] Translation of the full-length DNA sequences for MaSp1 and MaSp2 produces highly repetitive, Gly-Ala-rich large molecular weight proteins.[10] Protein architectures of MaSp1 and MaSp2 include non-repetitive, conserved N- and C-terminal domain (NTD and CTD) and internal block repeats consisting of about 35-40 amino acids. The NTD and CTD have been implicated in controlling the silk assembly process, and specifically, the CTD has been reported to mediate solubility and storage in the ampulla of the MA gland.[11-13] The repetitive regions of the spidroins contribute to the strength and extensibility of the fibers. Specifically, the repetitive segments of MaSp1 are dominated by GGX (X = A, Q, or Y), GA and poly-A (4-10 consecutive A), whereas the MaSp2 amino acid sequence includes GPX (X = G or S), QQ, GGX (X is typically A), GSG, and poly-A (3-9 consecutive A).[10] During storage of MaSp1 and MaSp2 in the ampulla, protein chain regions corresponding to the poly-A blocks

facilitate beta-sheet and nanocrystal formation, a structure that has been attributed to the high tensile strength of extruded fibers.[6, 14-16] Given the superior mechanical properties of spider silk, but impracticability of farming arachnids on a large-scale format, there has been a concerted effort by scientists across the globe to manufacture synthetic spider silk for commercial applications. In order to achieve mass production of artificial fibers, investigators have focused their attention on two areas: 1) development of heterologous expression systems capable of production of vast quantities of recombinant spidroins and 2) creation of spinning methodologies that recapitulate the extrusion process of the natural constituents of the major ampulla spinning dope.[17-19]

Transcriptomic analyses of the cob-weaver spider *Latrodectus hesperus* (black widow) have shed light on highly expressed transcripts within silk-producing glands.[20] MaSp1 and MaSp2, along with 30 other gene transcripts, were more abundantly expressed in MA glands relative to cephalothoraxes, suggesting additional potential candidates for involvement in spider silk assembly of dragline silk or as constituents of the fibers themselves.[20, 21] Further proteomic studies using mass spectrometry have led to the identification of non-MaSp proteins in spider silk.[22, 23] These non-MaSp proteins include a family of low molecular weight cysteine-rich proteins referred to as CRPs. In black widow spiders, these encompass CRP1, CRP2, CRP3, CRP4 and CRP5.[23] Three of the CRPs, CRP1, CRP2, and CRP4, have been demonstrated to be constituents of the MA gland, spinning dope, and dragline silk fibers.[24] The presence of CRP1, CRP2, and CRP4 in these structures suggests an important biological role in assembly of dragline silk and/or fiber mechanics. Computational modeling of the primary sequences of CRP family members suggests these proteins fold into a cystine slipknot.[23] Proteins with cystine slipknot motifs have been identified as the strongest force clamps known in the protein world,

suggesting that the CRPs potentially serve a mechanical role in dragline silk.[25] To date, the majority of studies have concentrated on the biophysical and mechanical properties of spidroins and little, if any, attention has been focused on CRP family members, despite their established presence in natural dragline silk.

In this study, we expressed and purified recombinant CRP1 and CRP4 from bacteria to characterize their secondary structure and investigated the impacts of pH and temperature on CRP folding using circular dichroism. My findings represent the first report to characterize the solution state structure of CRP family members, demonstrating that these proteins display high thermostability and structural conformations that are resistant to acidic pH environments. Elucidation of the in-solution structures of the CRPs is important to advance our understanding of the molecular mechanisms that govern spidroin assembly, silk extrusion and fiber performance. Our discoveries also provide important biochemical data that can be applied and integrated into synthetic spider silk production.

Chapter 2: Materials and Methods

Plasmid Construction and Expression of Recombinant Proteins in Bacteria

CRP1 and CRP4 cDNAs were amplified by PCR and inserted into the prokaryotic pET19b-10xHis-SUMO expression vector. CRP cDNAs were inserted in-frame into the expression vector, allowing for creation of N-terminal 10xHis-SUMO tags on the CRPs. The pET19b-10xHis-SUMO-CRP1 and pET19b-10xHis-SUMO-CRP4 expression vectors were transformed into BL21(DE3)pLysS competent cells (Promega). Transformants were grown overnight and then used for protein inductions with 1mM Isopropyl-beta-D-thiogalactoside (IPTG) (GoldBio). After 4 hrs of induction at 37°C, cells were pelleted and lysed in 1x FastBreak Cell Lysis Reagent (50 mM NaH_2PO_4, 300 mM NaCl, and 10 mM imidazole, pH 8.0) according to the manufacturer's instructions (Promega). Cell extracts were sonicated for 2 min in 10 sec pulses to ensure complete lysis. Following sonication, the solutions were clarified by centrifugation at 16,000 x g for 10 min at 4°C. Supernatants were filtered through a 0.2 micron filter and then used for affinity chromatography purification.

Purification of CRP1 and CRP4 using Ni-NTA Agarose Affinity and Removal of the SUMO-tag

Recombinant 10xHis-SUMO-CRP1 and CRP4 fusion proteins were purified using Ni-NTA affinity column chromatography. Bacterial supernatants were mixed with Ni-NTA Superflow Agarose resin (Qiagen), followed by incubation at 4°C for 1 hr with gentle agitation. After binding, the resin was washed 2 times with wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, pH 8.0), followed by detachment of the recombinant CRP fusion proteins

from the beads with elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole, pH 8.0). Following elution, imidazole was removed from the purified samples using Spectra/Por® Biotech cellulose dialysis tubing with a molecular weight cut-off of 500-1000 Da (Sigma-Aldrich) against a buffer (50 mM NaH_2PO_4, 300 mM NaCl, pH 8.0). 10xHis-tag-SUMO tags were removed from recombinant fusion proteins by treatment with purified His-tagged Ubiquitin-Like-Specific Protease 1 (ULP1) at 25°C for 1 hr. Following digestion, His-tagged ULP1 and 10xHis-SUMO segments were removed by re-applying digests to Ni-NTA resin. CRPs (lacking 10xHis-SUMO) failed to interact with the resin and were present in the flow-through fraction. Then, non-tagged CRPs were concentrated and de-salted using Amicon® Ultra-15 Centrifugal Filter Units with an Ultracel-3K membrane (EMD Millipore) using a series of water washes. Protein concentrations were determined using a BCA protein assay kit according to the manufacturer's instructions (Pierce Chemical Company).

SDS-PAGE and Western Blot Analysis

Samples were size fractionated using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) at 200V for 30 min. For protein detection, molecules were visualized with the ProteoSilver™ Silver Stain Kit according to the manufacturer's instructions (Sigma-Aldrich). Images were captured on a ChemiDoc™ XRS+ system (Bio-Rad Laboratories). For western blot analysis, proteins were transferred onto a 0.2 μm nitrocellulose membrane with a Pierce™ Power Blot Cassette (Thermo Fisher Scientific) for 4 min at 25 Volts and 1.3 Amperes. Prior to incubation with primary antibody, unbound sites on the membrane were blocked with 5% (w/v) non-fat dry milk dissolved in PBST (Phosphate Buffered Saline supplemented with 0.05% Tween-20). Antigens were detected using an anti-His mouse

monoclonal antibody (Santa Cruz Biotechnology, Inc.). Antigen-antibody complexes were detected using a secondary goat anti-mouse antibody conjugated with horseradish peroxidase (goat anti-mouse HRP antibody). Protein-antibody interactions were visualized using chemiluminescence with a ChemiDoc™ XRS+ system using the Luminata Crescendo Reagent (Bio-Rad Laboratories).

Tryptic and GluC Digests and Protein Preparation

Purified CRP1 or CRP4 (0.15mg/mL) were denatured in 100mM sodium phosphate and 6M Urea. Proteins were reduced with dithiothreitol (DTT) and alkylated with iodoacetamide (IAA). After reducing the urea concentration to 1M, CRP1 or CRP4 were digested with Trypsin (1:100) and GluC (1:200) (Pierce Chemical Company) for 18 hr at 37°C. Peptide clean-up was achieved by Pierce C18 Spin Columns, following the manufacturer's instructions (Thermo Scientific Pierce).

NanoLC Orbitrap Fusion™ Tribrid™ Mass Spectrometry and Protein Confirmation

Tryptic and GluC digests of purified recombinant CRP1 and CRP4 (minus 10xHis-SUMO tags) were analyzed by LC-MS/MS on a Thermo Scientific™ Orbitrap Fusion™ Tribrid Mass™ Spectrometer equipped with nanoHPLC coupled with an EASY-Spray™ ion source. Digests (5 μL or 0.25 μg) were injected onto a 20 mm x 75 μm C18 trap column (Thermo Scientific™) and run at a flow rate of 0.3 μL/min. The trap column was positioned in-line with the analytical EASY-Spray LC column (150 mm x 75 μm and particle size of 3 μm) for mass spectrometry analysis. Linear gradients of 1-50% solvent B over 33 min at 0.3 μL/min flow rate, followed by a steeper gradient from 50-99%, were used for peptide elution. Solvent B was held

at 99% for 12 min to wash the column and returned to 1% solvent B. Before injection of the next sample, however, the column was washed by injecting water and running HCD fragmentation with the same chromatography method used for sample analysis. Solvents A and B consisted of 0.1% formic acid (aqueous) and 100% acetonitrile, respectively. The ionspray voltage was 1700V, and interface heater was set at 300°C. Full scan product ion data (MS^1) were acquired by implementing a data dependent acquisition mode using Xcalibur 4.0 software. Full scans were collected with the orbitrap mass analyzer using an AGC target of 1 million, with a maximum inject time of 100 ms and resolution setting of 120,000 (quadrupole isolation window between *m/z* 300-1500 Da). Peptide ions observed with orbitrap-MS scans exceeding a threshold of 100,000 and charge state of +2 to +8 were set to trigger acquisition of product ion fragmentation (quadrupole isolation window: *m/z* 1 Da, AGC target of 100,000, and injection time of 200 ms) where MS/MS spectra of the resultant 20 most intense ions were collected by HCD. The LC-MS/MS data were analyzed with Proteome Discoverer 2.1 (Thermo Fisher Scientific). The database search included an annotated Uniprot database for *L. hesperus* (downloaded March 17, 2016) containing all SwissProt and Trembl entries. In addition, a *L. hesperus* transcriptome was downloaded on July 29, 2016 from NCBI (BioProject PRJNA242358) and converted to an unannotated proteome fasta file using tools from galaxy.org.[21] Briefly, getORF was used to retrieve open reading frames (ORFs), and then ORFs were converted to polypeptides using an *in silico* translational tool. A third custom database was manually curated with known protein sequences that were not present in the Uniprot database, namely CRP1 and CRP4. Precursor ions with a mass between 550 Da and 8500 Da were searched against the Uniprot database, using SEQUEST HT with a precursor mass tolerance of 500 ppm and fragment mass tolerances of 0.6 Da for HCD. Two fixed modifications were

allowed, carbamidomethylation of cysteine and N-terminal acetylation, as well as the dynamic modification for oxidation of methionine.

Circular Dichroism Spectroscopy and Dichroweb

Purified protein samples were analyzed using a J-810 Spectropolarimeter with an MPTC-490S Auto Peltier 6-cell changer attached to a water pump temperature controller (Jasco Inc.) using a quartz cuvette (Hellma, Sigma-Aldrich) of 0.1 cm path length. Data were collected using the SpectraManager software (Jasco Inc.), where measurements were taken with 3 accumulations from 260 to 190 nm with a data pitch of 0.5 nm and a scanning speed of 50 nm/min. Baseline measurements were taken using 10 mM potassium phosphate buffer at the desired pH. Data was exported to Microsoft Excel as xy-coordinates and graphed on Plot2. Thermal denaturation spectra were monitored at 208 nm from 20°C to 90°C at intervals of 0.5°C using a heating ramp rate of 0.5°C/min and temperature tolerance of 0.1°C. CRP1 and CRP4 were buffered against 10 mM potassium phosphate (pH of 5.8, 6.8 or 7.0) and samples contained 50 μM protein concentration. CD data were analyzed using Dichroweb to quantify the effects of pH on the CRP proteins. The Microsoft Excel sheet containing millidegrees data from 260 to 190 nm at a wavelength step of 0.5 nm was converted into a tab delimited text file and uploaded to Dichroweb under the "Free (with preview)" file format. Data were analyzed using the CDSSTR and CONTIN analysis programs coupled with the SMP180 reference set optimized for 190-240 nm. Structural values from both analysis programs were averaged. The Mean Residue Weight (MRW) was calculated for both CRPs and the cell path length used was 0.1cm. Results were checked to confirm that the theoretical curves produced by the software matched the experimental curves.

Chapter 3: Results

Purification of Recombinant CRPs

The CRP1 and CRP4 cDNAs encode low molecular weight proteins that are cysteine-rich, each containing 12 conserved cysteine residues (Figure 1). MS/MS studies have demonstrated CRP1 and CRP4 are constituents of dragline silk fibers, highlighting the importance of elucidation of their function, especially as laboratories across the globe are endeavoring to manufacture protein blends to spin synthetic silk for commercialization.

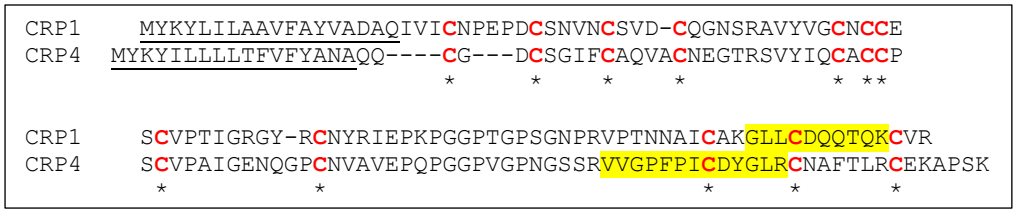

Figure 1: Amino acid sequence alignment of black widow CRP1 and CRP4. Single letter designations represent amino acid residues, while the underlined region represents the secretion signal predicted by SignalP 4.1 (DTU Bioinformatics). Red letters or asterisks represent conserved Cys (C) residues. Highlighted yellow regions correspond to select peptides identified by MS/MS analysis (see Figure 3).

To study the physical and chemical properties of CRP family members, 10xHis-tagged SUMO-CRP1 and CRP4 fusion proteins were expressed and purified from crude bacterial lysate using Ni-NTA affinity chromatography (Figure 2A-B). After collection of the flow-through (FT), wash (W), and elution (E) fractions, protein samples were size-fractionated using SDS-PAGE analysis, followed by visualization using silver staining (Figures 2A-B, lanes 2-6). Analyses of CRP1 and CRP4 FT fractions revealed dark smearing patterns, likely due to the presence of endogenous bacterial proteins unable to bind the nickel resin (Figure 2A-B, lane 2,

respectively). Wash fraction one (W1) contained similar, but less prominent smearing patterns, while the last wash fraction (W4) appeared to lack protein, signifying that non-specific bacterial proteins were removed from the resin (Figure 2A-B, lanes 3-4, respectively). The first two elution fractions (E1-2) revealed distinct protein bands migrating at approximately 25 kDa (Figure 2A-B, 5-6, respectively). These experimental masses were consistent with predicted molecular masses of the 10xHis-SUMO-CRP fusion proteins (SUMO and CRP1 or 4 predicted molecular masses are approximately 15- and 10-kDa, respectively). Western blot analysis using an anti-His monoclonal antibody confirmed that these species corresponded to SUMO-CRP1 and SUMO-CRP4 fusion proteins (data not shown).

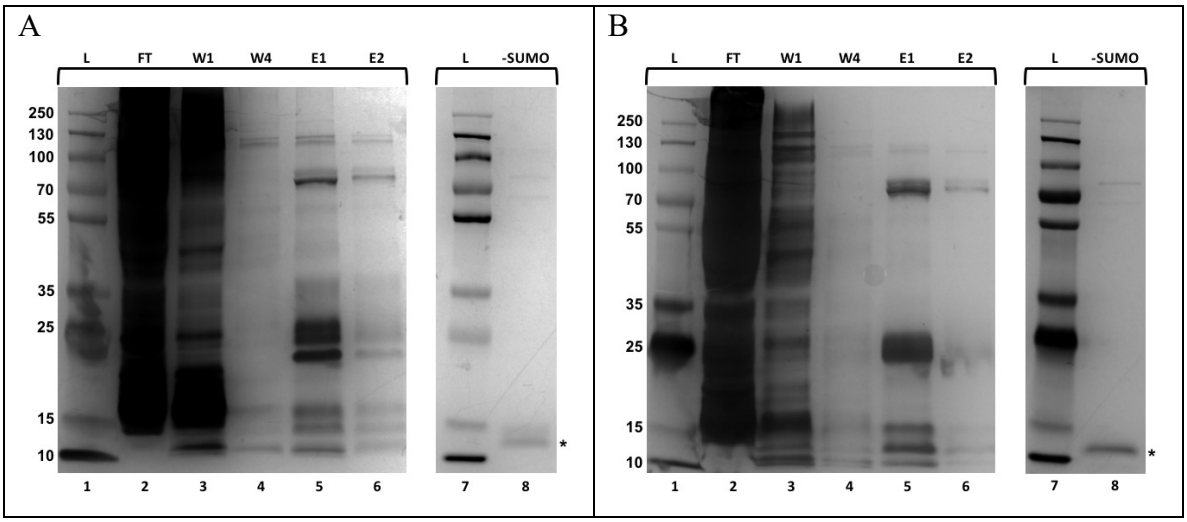

Figure 2: Isolation of recombinant black widow CRP1 (A) and CRP4 (B) from *E. coli* crude extract. SDS-PAGE analysis of purified CRP1 and CRP4 after affinity chromatography using a Ni-NTA agarose resin. Ubiquitin-like protease (ULP1) was used to remove the 10xHis-SUMO tags from purified 10xHis-SUMO-CRP1 and 10xHis-SUMO-CRP4. Proteins were visualized using silver staining. FT = flow through; W1 and W4 = wash 1 and 4; E = elution 1 and 2. L = protein ladder in kDa.

Because the 10xHis-SUMO appendage could potentially influence secondary structural characteristics of the CRPs, we employed a strategy to remove the 10xHis-SUMO tag from the CRPs. Purified 10xHis-SUMO-CRP1 and 10xHis-SUMO-CRP4 were treated with recombinant Ubiquitin-like-specific protease 1 (ULP1) to remove the 10xHis-SUMO segment. After incubation of 10xHis-SUMO-CRP1 or CRP4 with ULP1, His-tagged ULP1, 10xHis-SUMO and the CRP components were reapplied to a nickel-NTA agarose column, which allowed for capturing of the His-tagged ULP1 and 10xHis-SUMO region. Analysis of the ULP1-cleaved flow-through samples for CRP1 and CRP4 by SDS-PAGE analysis, followed by silver staining, detected bands migrating at the expected mass range for CRP1 and CRP4 (Figure 2A-B, lanes 8 [-SUMO], respectively). To confirm that these bands represented CRP1 and CRP4, we performed in-solution tryptic digestions of the flow-through samples. Tryptic digests were then analyzed by an Orbitrap Fusion™ Tribrid™ mass spectrometer equipped with nano-HPLC. MS/MS data were acquired using high collision dissociation fragmentation and orbitrap detection (HCD-OT). Two product ion spectra from peptides 1190.58 and 1492.76 ($[M+2H]^{2+}$) corresponded to peptide sequences GLLCDQQTQK and VVGPFPICDYGLR derived from CRP1 and CRP4, respectively (Figure 3A-B; also see Figure 1). MS/MS analyses led to 100% and 94% sequence coverage for CRP1 and CRP4, respectively. No SUMO or His-tagged ULP1 were detected in the flow through fraction, indicating that purified CRP1 and CRP4 completely lacked the SUMO tag.

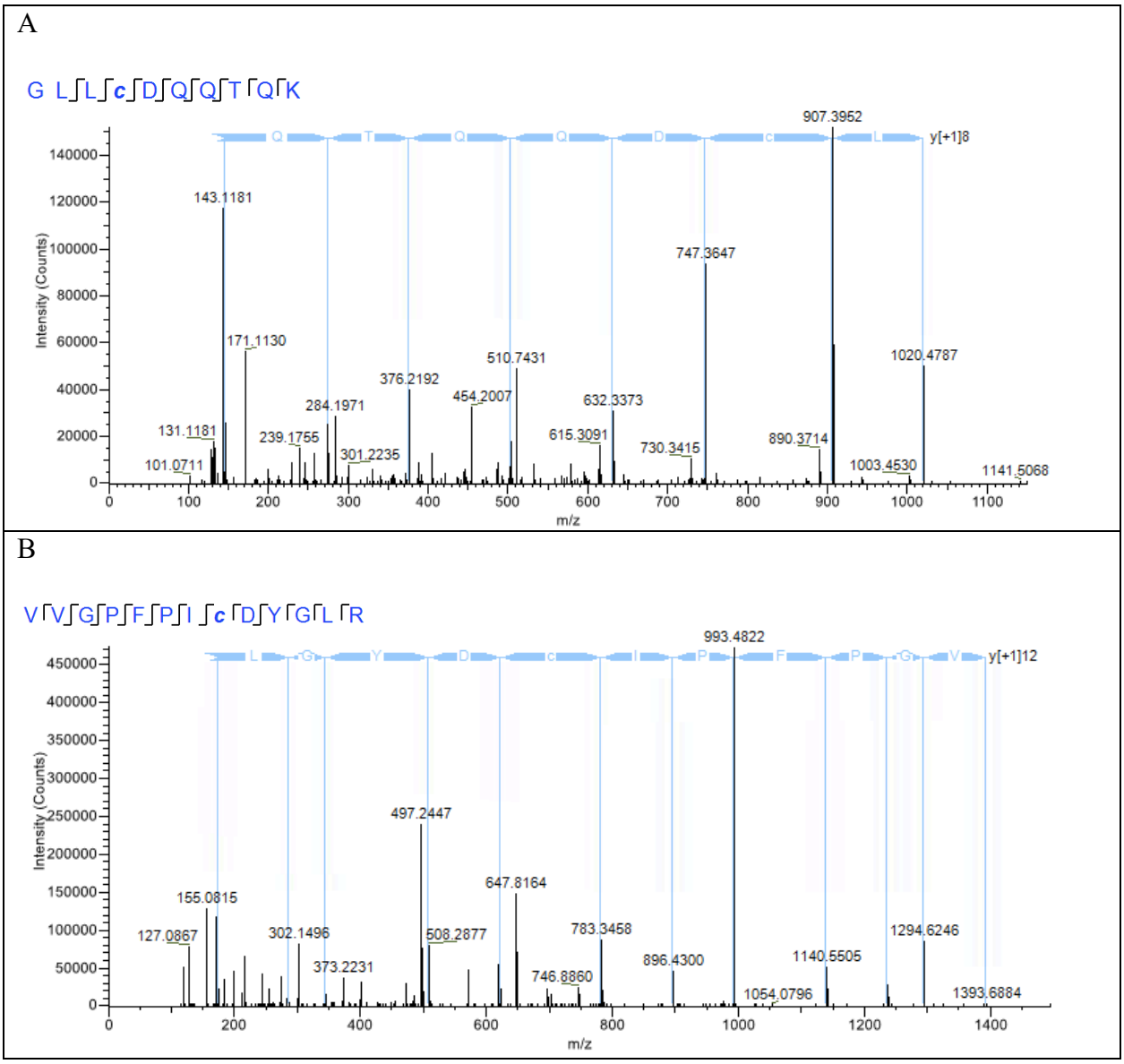

Figure 3: MS/MS spectra of purified CRP proteins after in-solution tryptic digestion using an Orbitrap Fusion™ Tribrid™ mass spectrometer confirm that purified proteins correspond to CRP1 and CRP4 lacking 10xHis-SUMO tags. (A) CRP1; (B) CRP4.

Secondary Structural Analysis of CRPs Under Different pH Conditions

After MS/MS analyses confirmed the successful purification of CRP1 and CRP4 by chromatography, the recombinant proteins were subjected to circular dichroism (CD)

spectroscopy to characterize secondary structural characteristics. Structural conformations of CRP1 and CRP4 were analyzed under 3 distinct pH-buffered potassium phosphate solutions, with pH of 5.8, 6.8, and 7.0 (Figure 4A-B). These pH conditions fall within hydrogen ion concentrations reported within the ampulla of the MA gland and spinning duct of spiders.[26] Secondary structures were investigated by CD from 190-260 nm at 25°C, a temperature that closely corresponds to physiological conditions of the silk-producing glands of black widow spiders.[26] At pH 7, the most abundant secondary structure of CRP1 was random coil conformation, with lower levels of beta sheet, beta turns and alpha helical structure (Figure 4A; Table 1). These percentages corresponded to approximately 42% random coil, 35% beta sheet, 12% turns, and 10% alpha helices (Table 1). Under acidic or slightly acidic conditions of pH = 5.8 or 6.8, respectively, the secondary structure of CRP1 was not altered significantly (Figure 4A; Table 1). Analysis of the CRP4 CD spectra at pH 7 revealed approximately 45% random coil, 32% beta-sheet, 11% beta-turns, and 11% alpha helical structure (Figure 4B; Table 2). Similar conformations were observed at pH 5.8 and 6.8 (Figure 4B and Table 2). Comparatively speaking, the secondary structures of CRP1 and CRP4 at pH 5.8 to 7 generated similar profiles (compare Figure 4A-B, Tables 1-2). Consistent with these experimental observations, the secondary structures of CRP1 and CRP4 closely align to predictions made by the NetSurfP 2.0 modeling algorithm (Figure 4C-D). Collectively, these findings support that CRP1 and CRP4 have substantial amounts of random coil conformation, followed by slightly lower levels of beta-sheet structure. Overall, these studies also reveal that there are little, if any, changes to secondary structure across a pH range of 5.8 to 7.0.

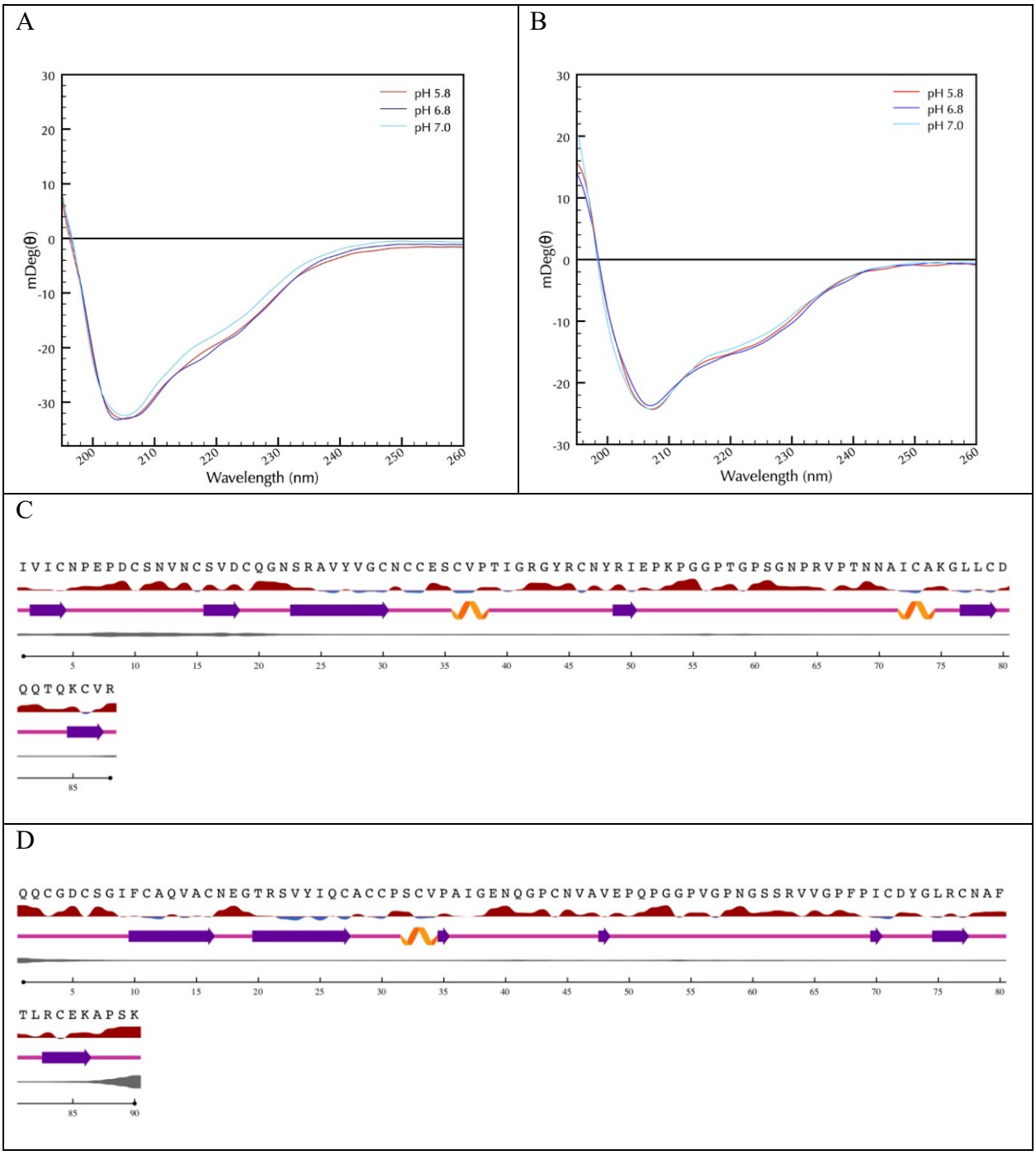

Figure 4: CD analyses of purified CRP1 or CRP4 under different pH conditions reveal large amounts of random coil and beta sheet conformation. Recombinant CRP1 or CRP4 were placed into solutions of different pH ranging from 5.8 to 7.0 and their secondary structural properties were investigated at 25°C. A) CRP1; B) CRP4. Predicted secondary structure produced by NetSurfP-2.0, depicted as a 3-state protein secondary structure model. Orange wave-like structure denotes helical regions, purple arrows represent beta sheet structure, and pink straight lines correspond to random coil. Blue and red mounds depict buried and exposed regions, respectively. C) CRP1; D) CRP4.

Table 1: Secondary structure of CRP1 broken down into the four main types of secondary structure identified by DichroWeb.

Percentage of each Characteristic at 25°C for CRP1				
	Random Coil	**Beta sheet**	**Turns**	**Helix**
pH 5.8	45.6	31.35	11.95	10.6
pH 6.8	45.1	31.1	11.95	11.4
pH 7.0	41.85	35.25	11.75	10.15
Predicted*	43.5	30.7	18.2	7.9

*Predicted percentages were calculated from the 8-state protein secondary structure (SS_8) predicted model produced by NetSurfP-2.0, with Bend collapsed into Random Coil structure.

Table 2: Secondary Structure of CRP4 broken down into the four main types of secondary structure identified by DichroWeb.

Percentage of each characteristic at 25°C for CRP4				
	Random Coil	**Beta sheet**	**Turns**	**Helix**
pH 5.8	44.95	33	11.25	10.3
pH 6.8	42.85	34.3	11.65	10.65
pH 7.0	45	32.35	11.15	11.05
Predicted*	45.5	33.3	17.8	3.3

*Predicted percentages were calculated from the 8-state protein secondary structure (SS_8) predicted model produced by NetSurfP-2.0, with Bend collapsed into Random Coil structure.

Secondary Structural Analysis of CRPs Under Different Temperature and pH Conditions

After investigating the impact of pH on the secondary structure of CRP1 and CRP4 at 25°C, we examined the influence of pH under different temperature conditions to explore thermal unfolding of the CRPs. Using CD, we generated thermal denaturation curves from 20°C to 90°C at 208 nm, a wavelength that corresponded to the minima of the secondary curves of CRP1 and CRP4 (Figure 4A-B). Within the investigated pH range of 5.8 to 7.0, the far-UV spectrum of native CRP1 and CRP4 revealed proteins that are extremely resistant to unfolding at elevated temperatures, indicating enormous conformational stability up to 90°C (Figure 5A-B).

In fact, at 90°C the CD data suggests that the CRPs have not yet approached their T_m. Attempts to perform a Protein Thermal Shift assay using GloMelt™ (Biotium) to detect protein unfolding with fluorescence were unsuccessful (data not shown), an observation that supports that substantial amounts of hydrophobic residues are pre-exposed to the solvent while the CRPs are present in their native conformational state. Thus, these studies support that CRP1 and CRP4, in their native states, display both conformational stability from pH 5.8 to 7.0 and thermal tolerance from 20 to 90°C.

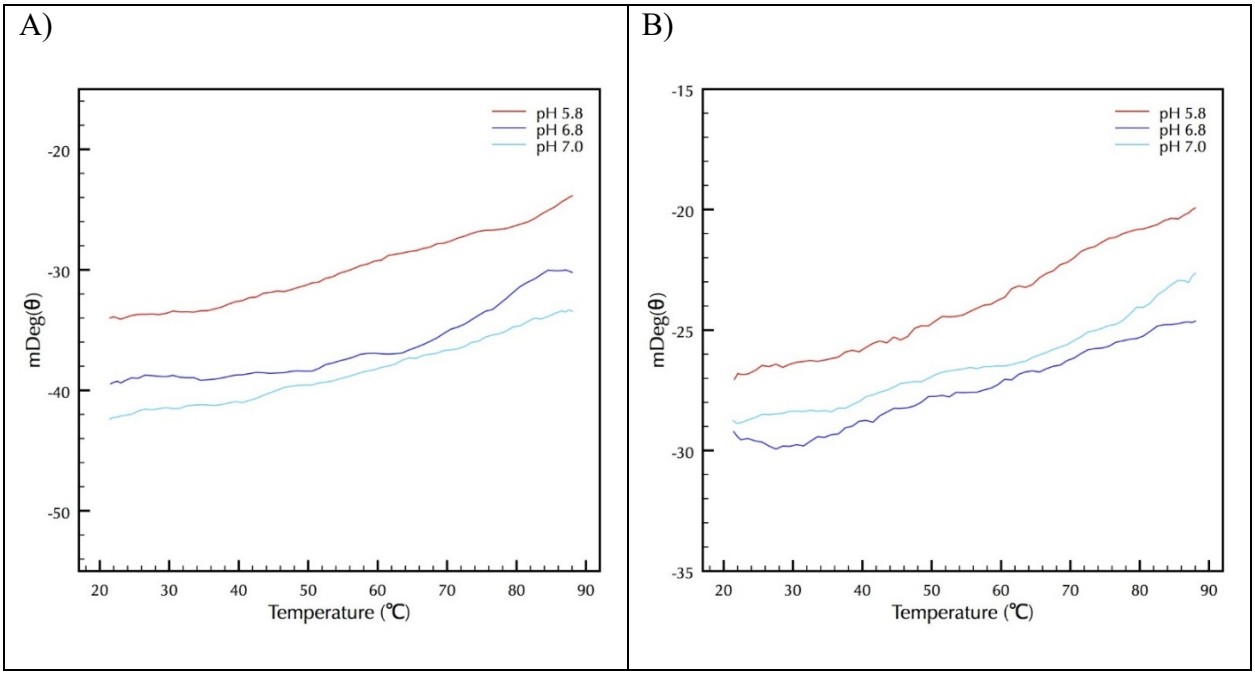

Figure 5: Thermal denaturation studies of CRPs under different pH environments reveal proteins with high thermal stability. The molar ellipticity measured at 208 nm was monitored continuously in the temperature range from 20 to 90°C by CD. Recombinant proteins were buffered at different pH of 5.8, 6.8, and 7.0. A) CRP1; B) CRP4.

Chapter 4: Discussion

The molecular mechanisms that govern spidroin assembly and identification of constituents involved in the extrusion process are fundamental aspects that must be elucidated before synthetic silks can be successfully spun for commercialization. In our studies, we provide new insight regarding structural features of two CRP family members, demonstrating that CRP1 and CRP4 have secondary structures that are largely random coil and beta sheet. Both recombinant proteins were expressed and purified as 10xHis-SUMO-fusion proteins, followed by successful removal of the 10xHis-SUMO tags by treatment with ULP1. Using purified CRP1 and CRP4 recombinant proteins, we show that CRPs experience little, if any, structural transitions in response to chemical changes, specifically hydrogen ion concentration that are in the acidic range. Changes in pH have been reported to play a significant role in regulating spidroin assembly. Specifically, studies in orb-weavers have demonstrated that the spinning dope pH in the ampulla is 7.0, but drops to approximately 6.3 in the spinning duct.[27, 28] This increase in acidity has been proposed to change the charge-state of specific R-groups on amino acids within spidroin protein chains, resulting in conformational changes, an event that helps trigger aggregation, assembly, and alignment of spidroins.[11, 29, 30] Both the N- and C-terminal domains (NTD and CTD) of the MaSps have been proposed to be important modulators of this process.[11, 31] In fact, it has been reported that monomeric MaSp NTDs can dimerize when the pH changes from neutral to acidic conditions.[11] This pH-regulated process has been implicated in triggering spidroin aggregation prior to extrusion.[32] Although proteomic studies have confirmed the presence of CRPs in cob weaver spider silk-producing glands, specifically CRP1, CRP2, and CRP4 in the MA gland, spinning dope, and dragline silk fiber, the precise biological role of the CRPs is at present unknown.[24] Whether CRPs physically interact with spidroins is currently

unclear; however, mechanistically speaking, our data support that acidification is unlikely to trigger CRP-conformational changes during extrusion of these proteins from the ampulla through the spinning duct, a central chemical event that induces spidroin assembly. Based on quantitative real-time PCR studies and proteomic analyses, we have previously demonstrated that CRP4 is transcribed and translated in the MA gland at higher levels than CRP1, and then subsequently secreted into the dope and fibers. This suggests that despite their similarities in amino acid sequences, CRP1 and CRP4 may have distinct biological functions.[23, 24] However, based on our CD studies performed under different pH and temperature conditions, the structural characteristics of CRP1 and CRP4 appear similar. Thus, more biochemical characterization will be necessary to elucidate the functional difference between CRP1 and CRP4, as well as more future studies to explore the structure of CRP2, CRP3, and CRP5.

Further studies should be completed to explicitly define the disulfide bond linkages within the CRPs. This could allow for the confirmation of the cystine slipknot motif and further evidence to support that these proteins may function as a molecular clamp in dragline silk. Additionally, site directed mutagenesis should be done to alter cysteine residues of the CRPs followed by thermal stabilities studies. These studies would give further insight into the specific bonds of the CRPs that contribute to the high thermal stability of the recombinant proteins.

Chapter 5: Conclusion

Molecular modeling and sequence alignments with other proteins in the protein database suggest CRP members contain a cystine-slipknot motif.[23] This structure is present within a fairly large number of secreted proteins and glycoproteins that belong to the TGFbeta and glycoprotein hormone (GPH) superfamily.[33, 34] The cystine knot motif consists of three intertwined disulfide bonds. This structure has been described as one of the strongest force clamps known in nature, supported by molecular dynamics simulations.[25] Our findings support natively folded CRPs contain high thermal stability, which is consistent with other cystine knot proteins, such as α-amylase inhibitor, who has a T_m over 100°C.[35] In addition, other cystine knot-containing proteins are known to lack hydrophobic cores.[35, 36] CRPs also lack hydrophobic cores, which is supported by our GloMelt™ thermal shift assay studies and NetSurfP-2.0 prediction program results. Moreover, far-UV CD spectra reported from other cystine knot proteins closely resemble the CD spectra of the CRPs, implying structural similarities with established cystine knot proteins.[36] From a structural perspective, the biophysical properties of CRPs are consistent with reported cystine slipknot proteins, further suggesting CRPs serve a mechanical role in black widow spider threads.

References

[1] Gosline, J. M., DeMont, M. E., Denny, M. W. (1986) The structure and properties of spider silk, *Endeavour 10*, 31-43.

[2] Moore, A. M., and Tran, K. (1999) Material properties of cobweb silk from the black widow spider *Latrodectus hesperus*, *Int J Biol Macromol 24*, 277-282.

[3] Foelix, R. (1996) *Biology of spiders*, Oxford University Press, New York.

[4] Guerette, P. A., Ginzinger, D. G., Weber, B. H., and Gosline, J. M. (1996) Silk properties determined by gland-specific expression of a spider fibroin gene family, *Science 272*, 112-115.

[5] Andersson, M., Holm, L., Ridderstrale, Y., Johansson, J., and Rising, A. (2013) Morphology and composition of the spider major ampullate gland and dragline silk, *Biomacromolecules 14*, 2945-2952.

[6] Vollrath, F., and Knight, D. P. (2001) Liquid crystalline spinning of spider silk, *Nature 410*, 541-548.

[7] Jin, H. J., and Kaplan, D. L. (2003) Mechanism of silk processing in insects and spiders, *Nature 424*, 1057-1061.

[8] Xu, M., and Lewis, R. V. (1990) Structure of a protein superfiber: spider dragline silk, *Proc. Natl. Acad. Sci. USA 87*, 7120-7124.

[9] Hinman, M. B., and Lewis, R. V. (1992) Isolation of a clone encoding a second dragline silk fibroin. *Nephila clavipes* dragline silk is a two-protein fiber, *J. Biol. Chem. 267*, 19320-19324.

[10] Ayoub, N. A., Garb, J. E., Tinghitella, R. M., Collin, M. A., and Hayashi, C. Y. (2007) Blueprint for a high-performance biomaterial: full-length spider dragline silk genes, *PLoS ONE 2*, e514.

[11] Hagn, F., Eisoldt, L., Hardy, J. G., Vendrely, C., Coles, M., Scheibel, T., and Kessler, H. (2010) A conserved spider silk domain acts as a molecular switch that controls fibre assembly, *Nature 465*, 239-242.

[12] Landreh, M., Askarieh, G., Nordling, K., Hedhammar, M., Rising, A., Casals, C., Astorga-Wells, J., Alvelius, G., Knight, S. D., Johansson, J., Jornvall, H., and Bergman, T. A pH-dependent dimer lock in spider silk protein, *J Mol Biol 404*, 328-336.

[13] Kronqvist, N., Otikovs, M., Chmyrov, V., Chen, G., Andersson, M., Nordling, K., Landreh, M., Sarr, M., Jornvall, H., Wennmalm, S., Widengren, J., Meng, Q., Rising, A., Otzen, D., Knight, S. D., Jaudzems, K., and Johansson, J. (2014) Sequential pH-driven dimerization and stabilization of the N-terminal domain enables rapid spider silk formation, *Nat Commun 5*, 3254.

[14] Trancik, J. E., Czernuszka, J. T., Bell, F. I., and Viney, C. (2006) Nanostructural features of a spider dragline silk as revealed by electron and X-ray diffraction studies, *Polymer 47*, 5633-5642.

[15] Holland, G. P., Jenkins, J. E., Creager, M. S., Lewis, R. V., and Yarger, J. L. (2008) Solid-state NMR investigation of major and minor ampullate spider silk in the native and hydrated states, *Biomacromolecules 9*, 651-657.

[16] Jenkins, J. E., Sampath, S., Butler, E., Kim, J., Henning, R. W., Holland, G. P., and Yarger, J. L. (2013) Characterizing the secondary protein structure of black widow

dragline silk using solid-state NMR and X-ray diffraction, *Biomacromolecules 14*, 3472-3483.

[17] Xu, J., Dong, Q., Yu, Y., Niu, B., Ji, D., Li, M., Huang, Y., Chen, X., and Tan, A. (2018) Mass spider silk production through targeted gene replacement in *Bombyx mori*, *Proc Natl Acad Sci U S A 115*, 8757-8762.

[18] Bowen, C. H., Dai, B., Sargent, C. J., Bai, W., Ladiwala, P., Feng, H., Huang, W., Kaplan, D. L., Galazka, J. M., and Zhang, F. (2018) Recombinant Spidroins Fully Replicate Primary Mechanical Properties of Natural Spider Silk, *Biomacromolecules 19*, 3853-3860.

[19] Rammensee, S., Slotta, U., Scheibel, T., and Bausch, A. R. (2008) Assembly mechanism of recombinant spider silk proteins, *Proc Natl Acad Sci U S A 105*, 6590-6595.

[20] Lane, A. K., Hayashi, C. Y., Whitworth, G. B., and Ayoub, N. A. (2013) Complex gene expression in the dragline silk producing glands of the Western black widow (*Latrodectus hesperus*), *BMC Genomics 14*, 846.

[21] Clarke, T. H., Garb, J. E., Hayashi, C. Y., Haney, R. A., Lancaster, A. K., Corbett, S., and Ayoub, N. A. (2014) Multi-tissue transcriptomics of the black widow spider reveals expansions, co-options, and functional processes of the silk gland gene toolkit, *BMC Genomics 15*, 365.

[22] Chaw, R. C., Correa-Garhwal, S. M., Clarke, T. H., Ayoub, N. A., and Hayashi, C. Y. (2015) Proteomic evidence for components of spider silk synthesis from black widow silk glands and fibers, *J Proteome Res 14*, 4223-4231.

[23] Pham, T., Chuang, T., Lin, A., Joo, H., Tsai, J., Crawford, T., Zhao, L., Hsia, Y., Williams, C., and Vierra, C. A. (2014) Dragline silk: a fiber assembled with low-molecular-weight cysteine-rich proteins, *Biomacromolecules 10,* 4073-4081.

[24] Larracas, C., Hekman, R., Dyrness, S., Arata, A., Williams, C., and Vierra, C. (2016) Comprehensive Proteomic Analysis of Spider Dragline Silk from Black Widows: A Recipe to Build Synthetic Silk Fibers, *International Journal of Molecular Sciences 17*, 1537-1553.

[25] Sikora, M., Sulkowska, J. I., and Cieplak, M. (2009) Mechanical strength of 17,134 model proteins and cysteine slipknots, *PLoS Comput Biol 5*, e1000547.

[26] Vollrath, F., Wen Hu, X., and Knight, D. P. (1998) Silk production in a spider involves acid bath treatment, *Proc. R. Soc. B 263*, 817-820.

[27] Dicko, C., Vollrath, F., and Kenney, J. M. (2004) Spider silk protein refolding is controlled by changing pH, *Biomacromolecules 5*, 704-710.

[28] Dicko, C., Kenney, J. M., Knight, D., and Vollrath, F. (2004) Transition to a beta-sheet-rich structure in spidroin *in vitro*: the effects of pH and cations, *Biochemistry 43*, 14080-14087.

[29] Askarieh, G., Hedhammar, M., Nordling, K., Saenz, A., Casals, C., Rising, A., Johansson, J., and Knight, S. D. (2010) Self-assembly of spider silk proteins is controlled by a pH-sensitive relay, *Nature 465*, 236-238.

[30] Bauer, J., Schaal, D., Eisoldt, L., Schweimer, K., Schwarzinger, S., and Scheibel, T. (2016) Acidic Residues Control the Dimerization of the N-terminal Domain of Black Widow Spiders' Major Ampullate Spidroin 1, *Sci Rep 6*, 34442.

[31] Hagn, F., Thamm, C., Scheibel, T., and Kessler, H. (2011) pH-dependent dimerization and salt-dependent stabilization of the N-terminal domain of spider dragline silk--implications for fiber formation, *Angew Chem Int Ed Engl 50*, 310-313.

[32] Schwarze, S., Zwettler, F. U., Johnson, C. M., and Neuweiler, H. (2013) The N-terminal domains of spider silk proteins assemble ultrafast and protected from charge screening, *Nat Commun 4*, 2815.

[33] Scheufler, C., Sebald, W., and Hulsmeyer, M. (1999) Crystal structure of human bone morphogenetic protein-2 at 2.7 A resolution, *J Mol Biol 287*, 103-115.

[34] Vitt, U. A., Hsu, S. Y., and Hsueh, A. J. (2001) Evolution and classification of cystine knot-containing hormones and related extracellular signaling molecules, *Mol Endocrinol 15*, 681-694.

[35] Muller, Y. A., Heiring, C., Misselwitz, R., Welfle, K., and Welfle, H. (2002) The cystine knot promotes folding and not thermodynamic stability in vascular endothelial growth factor, *J Biol Chem 277*, 43410-43416.

[36] Hoffmann, A., Funkner, A., Neumann, P., Juhnke, S., Walther, M., Schierhorn, A., Weininger, U., Balbach, J., Reuter, G., and Stubbs, M. T. (2008) Biophysical characterization of refolded Drosophila Spatzle, a cystine knot protein, reveals distinct properties of three isoforms, *J Biol Chem 283*, 32598-32609.